POINT ESTIMATION AND CONFIDENCE INTERVAL ESTIMATION FOR BINOMIAL AND MULTINOMIAL PARAMETERS

Ramesh Chandra

Union College
Schenectady, New York

31 December 1975

POINT ESTIMATION AND CONFIDENCE INTERVAL ESTIMATION FOR BINOMIAL AND MULTINOMIAL PARAMETERS

AES-7514

ADMINISTRATIVE AND ENGINEERING SYSTEMS MONOGRAPH

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AES-7514	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Point Estimation and Confidence Interval Estimation for Binomial and Multinomial Parameters		5. TYPE OF REPORT & PERIOD COVERED Final--Jan. 1, 1975 to Dec. 31, 1975 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ramesh Chandra.		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0583
9. PERFORMING ORGANIZATION NAME AND ADDRESS Institute of Administration & Management, Union College, Schenectady, N.Y. 12308		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-302
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of Navy Arlington, Va.		12. REPORT DATE Dec. 31, 1975 13. NUMBER OF PAGES<br~~22~~ 24
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Part of doctoral dissertation at Union College and University.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Point estimation, sequential estimation, sequential confidence intervals, binomial, multinomial populations, Bayesian estimates.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study deals with the estimation of parameter(s) of binomial or multinomial distribution using the data available at the termination of a sequential experiment.

ia

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AES-7514

POINT ESTIMATION AND CONFIDENCE INTERVAL ESTIMATION FOR BINOMIAL AND MULTINOMIAL PARAMETERS

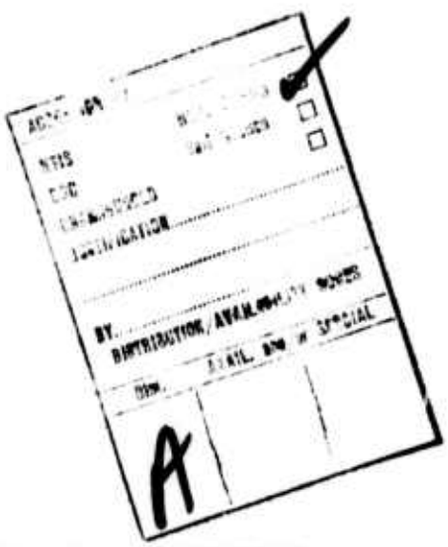
by

Ramesh Chandra

December 1975

Prepared with partial support
of the Office of Naval Research
Contract N00014-75-C-0583

L.A. Aroian, Project Director



UNION COLLEGE AND UNIVERSITY
Institute of Administration and Management
Schenectady, New York 12308

POINT ESTIMATION AND CONFIDENCE INTERVAL ESTIMATION FOR BINOMIAL AND MULTINOMIAL PARAMETERS

1. INTRODUCTION

After the completion of experimentation (for testing k hypotheses concerning binomial parameter p (R. Chandra, 1975-A), selection of the most probable category of a multinomial population (R. Chandra, 1975-B), and ranking of k binomial populations (R. Chandra, 1975-C)), the experimenter might wish to obtain some estimates of the value(s) of the unknown parameter(s) occurring in the probability distribution of the population(s). In this paper our basic theme is the estimation of parameter(s) of binomial and multinomial distributions from an "effectively small" sample. We suppose that a sample is already available to be used by itself for the estimation. By "effectively small" we mean that a sample is small for the purposes of estimating the probabilities, since the experiment was primarily designed for the separation/selection/ranking rather than for the estimation. The estimation, we assume, is the afterthought rather than the main objective of the experimenter.

If the primary objective of the experimenter is to obtain, with as few observations as possible, a confidence interval for an unknown parameter, having preassigned confidence coefficient

and preassigned width, or to satisfy some other conditions, he would need altogether a different stopping rule for the sequential experimentation. DeGroot (1959) has shown that a single sampling plan and inverse binomial plan are the only ones that admit an estimator that is efficient at all values of p of a binomial distribution. Bhat and Kulkarni (1966) have generalized the above results for multinomial distribution with the similar conclusion. Thus it is clear that if estimation of parameter(s) is the primary aim of the experimenter, he would choose one of the above two procedures. For further reference on sequential estimation one can refer to Johnson (1961), Wetherill (1966), and Goss (1974), among others. Our concern here is the development of an exact method of point and interval estimation for the parameter(s) once we have terminated our experiment for hypothesis testing/selection/ranking. We base our estimation solely upon the sample available on the completion of our experimentation.

Anscombe (1953, 1954) has made a number of studies of the use of the data obtained with sequential sampling procedures in the calculation of estimates of the parameter values. He has concluded that provided the sample size is not too small fixed sample size methods of estimation (which is generally maximum likelihood procedure) can be used even though the samples were obtained sequentially. Johnson (1961, p. 361) rightly notes "...these estimates would be no more accurate than would estimates based on genuine fixed size sample of the same size. In fact, owing to the essentially approximative ("large size") nature of

Anscombe's results, there would usually be rather more uncertainty as to the true accuracy of estimates based on sequentially obtained sample." Armitage has said "...it appears to be true that sequential estimation is no more efficient than fixed size sample estimation" (see discussion of Anscombe (1953)). Bearing the above facts in mind the experimenter must, therefore, accept that the efficiency of the estimates obtained from a sample collected for sequential experiments designed for purposes other than estimation would be limited.

Maximum likelihood procedure is the most often used method for obtaining estimates from one stage (or fixed size) sample. It is well known that maximum likelihood estimates based on a sequential sample are biased. Girshick et al. (1946) give an unbiased estimator for the binomial parameter. Their procedure has been extended for estimating the trinomial parameters by Muhamedhanova et al. (1966). The procedure can be generalized for $k>3$ (see Johnson et al. (1969)). Before introducing our method of estimation, we note the following reasons why we do not use either maximum likelihood procedure or methods based on Girshick et al. In the discussion that follows we consider the binomial distribution. The remarks that we make hold true for multinomial distribution also.

We take the maximum likelihood estimator (m.l.e.) first. Suppose we wish to estimate the probability p after we have ended the sequential experiment with i success(es) out of n trials. The m.l.e. is simply i/n. This is a perfectly good estimate

(ignoring the fact that it is biased) if i and j=n-i are both large and in this case there is no objection to m.l.e. However, since we are concerned with probability estimation from effectively small samples, let us consider the case where i=0. Then m.l.e. of p is 0. It is not easy to see how this value can be used for decision making. Unless n were fabulously large no one should make use of m.l.e when i=0. Confidence interval estimation suffers from a similar disadvantage when i=0 or it is very small; it has to include the point p=0. To assert that 0 is the lower bound for p is to assert nothing at all about the lower bound (for further discussion on this point, see Good (1965)).

Estimates based upon Girshick's method suffer the same shortcomings which we have pointed out above for m.l.e. Girshick's unbiased estimate $\hat{p}$ for p is

$$\hat{p} = w^*(i,n)/w(i,n)$$

where w(i,n) is the number of admissible paths from (0,0) to (i,n) and w*(i,n) is the number of admissible paths from (1,1). Clearly when i=0, w*(i,n)=0 and w(i,n)=1 and $\hat{p}$=0. Thus the objection noted above for m.l.e. for i=0 is applicable here also.

Noting the above difficulties with classical estimates, in what follows in this chapter we give a modern Bayesian procedure for estimation. We do not claim any optimality property for our procedure; we have attempted to provide a general and practical method of estimation after the termination of a

sequential experiment. In Section 2 we give the estimation procedure for binomial parameter p and in Section 3 we generalize our results of Section 2 for multinomial parameters. We derive the formulas with very general prior distribution of parameter(s) but we have restricted ourselves to the uniform prior distribution of the parameter(s) in our applications. Perk (1947), Johnson (1932), Good (1965), and Good et al. (1974), among others have discussed the pros and cons of using other priors vis-a-vis uniform prior. We, however, stick with uniform prior in our uses because this clearly represents best our state of ignorance about parameter(s) that we assume completely unknown at the start of experimentation. Important references on Bayesian estimation procedures include Wetherill (1966), DeGroot (1970), and Box and Tiao (1973) among others.

2. ESTIMATION OF BINOMIAL PARAMETER p

In this section we consider the estimation of the binomial parameter p. The Bayesian method that we describe and use for estimating p from the sample obtained at the termination of a sequential experiment is due to Goss (1974) and Schmee (1974). The method they have used is based on the assumption of a uniform prior distribution of parameter p between 0 and 1 (i.e. $G(p)=p$ and $dG(p)=g(p)=1$ for $0\leq p\leq 1$). In what follows, we first derive the formulas with a very general prior distribution of p and then give the formulas with a uniform prior. In Section 2.1 we derive a formula for point estimation. In Section 2.2 we derive

formulas for interval estimation. In Section 2.3 we give the results for a uniform prior. In Section 2.4 we give a numerical illustrative example.

2.1 Point Estimation

Let a sequential experiment terminate at the m^{th} stage with score i; i is the total number of successes associated with probability p. It is well known that i is a sufficient and transitive statistic.

Let g(p) be the prior density of p that is believed to be true, and $\ell(i,m)$ the likelihood function for the outcome. Then, according to Bayes' rule, the posterior density function, f(p), of distribution is given by:

$$f(p) = \frac{g(p)\ell(i,m)dp}{\int g(p)\ell(i,m)dp} \qquad (1)$$

where integration is over the whole range of p.

In Bayesian analysis the general form for g(p) is

$$g(p) = p^{a-1}(1-p)^{b-1};\ a,b\geq 0,\ 0\leq p\leq 1 \qquad (2)$$

(see Good (1965)). For experimental outcome (i,m) the likelihood function is given by:

$$\ell(i,m) = w(i,m)p^{i}(1-p)^{m-1} \qquad (3)$$

where w(i,m) is the total number of admissible paths from (0,0) to (i,m). (Note: $w(i,m) \neq \binom{m}{i}$, since for sequential procedures all the outcomes are not permutable. Fortunately,

however, w(i,m) does not present any problem, since it drops out as may be seen later.)

Substituting values of g(p) and $\ell(i,m)$ in (1) the posterior density becomes:

$$f(p) = \frac{p^{a-1}(1-p)^{b-1}w(i,m)p^{i}(1-p)^{m-i}dp}{\int_0^1 p^{a-1}(1-p)^{b-1}w(i,m)p^{i}(1-p)^{m-1}dp}$$

$$= \frac{p^{i+a-1}(1-p)^{m-i+b-1}dp}{\int_0^1 p^{i+a-1}(1-p)^{m-1+b-1}dp}$$

$$= \frac{p^{i+a-1}(1-p)^{m-i+b-1}dp}{B(i+a,\ m-i+b)} \qquad (4)$$

where B(x,y) is the complete beta function with paramters x and y; $B(x,y)=(x-1)!(y-1)!/(x+y-1)!$.

Let $\hat{p}$ be the point estimate of p, then

$$E(p) = \hat{p} = \int_0^1 p\ f(p)$$

$$= \frac{1}{B(i+a,m-i+b)} \cdot \int_0^1 p\ \ p^{i+a-1}(1-p)^{m-i+b-i}dr$$

$$= \frac{B(i+a+1,m-i+b)}{B(i+a,m-i+b)}$$

$$= \frac{(i+a)!(m-i+b-1)!}{(i+a+1+m-i+b-1)!} \cdot \frac{(i+a+m-i+b-1)!}{(i+a-1)!(m-i+b-1)!}.$$

or

$$\hat{p} = \frac{i+a}{m+a+b} \qquad (5)$$

2.2 Interval estimation

Let $\underline{p}$ be the lower bound on p at confidence level $1-\alpha$, $\alpha<1$ (i.e., $Pr(p<\underline{p})\leq\alpha$). It follows that

$$\int_0^{\underline{p}} f(p)\,dp = \alpha$$

or

$$\frac{1}{B(i+a,m-i+b)} \int_0^{\underline{p}} p^{i+a-1}(1-p)^{m-i+b-1}\,dp = \alpha$$

or

$$\frac{B_{\underline{p}}(i+a,m-i+b)}{B(i+a,m-i+b)} = \alpha \qquad (6)$$

where $B_{\underline{p}}(i+a,m-i+b)$ is incomplete beta function. Equation (6) may be solved for $\underline{p}$ by using any standard computer program. Also it is well known that $B_p(x,y)=B(x,y)\cdot E(x+y-1,x,p)$, where $E(n,r,p) = \sum_{x=r}^{n} p^x(1-p)^{n-x}$ and $0\le p\le 1$. Thus (6) reduces to:

$$E(m+a+b-1,\ i+a,\ \underline{p}) = \alpha \qquad (7)$$

The above equation may be solved for $\underline{p}$ using any standard table for cumulative binomial function (e.g. Harvard Table (1955)) with interpolation (if necessary).

Similarly, an upper bound $\bar{p}$ on p at confidence level β (i.e., $Pr(p\ge\bar{p})\le\beta$ or $Pr(p\le\bar{p})\ge 1-\beta$) may be computed by solving

$$\frac{B_{\bar{p}}(i+a,m-i+b)}{B(i+a,m-i+b)} = 1-\beta, \qquad (8)$$

or

$$E(m+a+b-1,i+a,\bar{p}) = 1-\beta \qquad (9)$$

The two-sided confidence interval at confidence level $1-\alpha-\beta$ (α from the lower side, and β from the upper side) may be obtained by solving (6) and (8) or (7) and (9) simultaneously.

2.3 Results with uniform prior

For uniform prior distribution of parameter p between

0 and 1 (inclusive) $g(p)=1$, i.e., $a=b=1$ in (2). Substituting the above values for a and b in (5)

$$\hat{p} = (i+1)/(m+2). \tag{10}$$

For fixed sample size test, the above result was first obtained by Laplace and is known as Laplace's law of succession (see Good (1965, p. 16).

The lower bound $\underline{p}$ on p at confidence level $1-\alpha$ may be obtained by solving one of the following two equations.

$$B_{\underline{p}}(i+1,m-i+1)/B(i+1,m-i+1) = \alpha \tag{11}$$

$$E(m+1,i+1,\underline{p}) = \alpha \tag{12}$$

Similarly, the upper bound $\bar{p}$ on p at confidence level β may be obtained by solving one of the following two equations:

$$B_{\bar{p}}(i+1,m-i+1)/B(i+1,m-i+1) = 1-\beta \tag{13}$$

$$E(m+1,i+1,\bar{p}) = 1-\beta \tag{14}$$

The two-sided bound $(\underline{p},\bar{p})$ on p at confidence level $1-\alpha-\beta(\alpha,\beta<1)$ may be obtained by solving (11) and (13) or (12) and (14) simultaneously.

2.4 Illustrative example

As an illustration of the explicit point and interval estimation, suppose our sequential experiment for ranking three

independent binomial populations A, B and C terminates at stage 3 with scores (successes) 3,2,0 for population A,B, and C respectively. It is desired to obtain:

a) Point estimates of population parameters p_A, p_B and p_C respectively;

b) One-sided bound on p_A, p_B, and p_C at confidence level 0.9;

c) Two-sided bound on p_A, p_B and p_C at confidence level 0.9 (equal tail).

Solution

a) From (10)

$$\hat{p}_A = \frac{3+1}{3+2} = 0.8,\ \hat{p}_B = \frac{2+1}{3+2} = 0.6,\ \hat{p}_C = \frac{1}{3+2} = 0.2$$

b) (i) Using (11) with $\alpha = 0.1$ (lower bounds)

$$\underline{p}_A = 0.562,\ \underline{p}_B = 0.320,\ \underline{p}_C = 0.025$$

(ii) Using (13) with $\beta = 0.1$ (upper bounds)

$$\overline{p}_A = 0.974,\ \overline{p}_B = 0.857,\ \overline{p}_C = 0.437$$

c) Using (11) and (13) with $\alpha = 0.05$ and $\beta = 0.05$, a 90% bounds on p_i (i=A,B,C) are as follows:

$$0.472 \le p_A \le 0.987$$

$$0.248 \le p_B \le 0.902$$

$$0.012 \le p_C \le 0.527$$

(Note: solutions for parts (b) and (c) were obtained using a computer program for evaluating inverse-incomplete beta function.)

3. ESTIMATION OF MULTINOMIAL PARAMETER $(p_1, p_2, \ldots, p_k)$

In this section we consider the estimation of multinomial parameter p_j associated with event E_j; $j=1,2,\ldots,k$. It is assumed that E_j's are exhaustive and mutually exclusive events such that $0<p_j<1$ and $\Sigma p_j=1$. The method that we describe below is the generalization of the one that we developed in Section 2. for the estimation of binomial parameter p. In what follows, we first derive the formulas with a very general prior for multinomial distribution and then give the results with uniform prior. In Section 3.1 we derive a formula for point estimation. In Section 3.2 we derive formulas for interval estimation. In Section 3.3 we give the results for uniform prior. In Section 3.4 we give a numerical illustrative example.

3.1 Point estimation

Let a sequential experiment terminate at the m^{th} stage with cumulative score n_j for event E_j; $0\leq n_j\leq m$, $\sum_{j=1}^{k} n_j=m$, $j=1,2,\ldots,k$. It is well known that any k-1 of the above n_j's constitute minimal sufficient statistics. It is also well known that a multinomial distribution of order k can be completely specified by any k-1 of the k p_j's. For convenience, we will consider the first k-1 of the k p_j's unless the contrary is mentioned.

Let $g(p_1 \ldots p_{k-1})$ be the prior density function of the distribution, and $\ell(n_1 \ldots n_k)$ be the likelihood function for the outcome. Then, according to Bayes' rule, the posterior

density, $f(p_1 \cdots p_{k-1})$, is given by

$$f(p_1 \cdots p_{k-1}) = \frac{g(p_1 \cdots p_{k-1}) \cdot \ell(n_1 \cdots n_k) \prod_{j=1}^{k-1} dp_j}{\int g(p_1 \cdots p_{k-1}) \cdot \ell(n_1 \cdots n_k) \prod_{j=1}^{k-1} dp_j} \qquad (15)$$

where integration is (k-1) fold over the whole range of p_j such that $\sum_{j=1}^{k} p_j = 1$.

In Bayesian analysis the general form for $g(p_1 \cdots p_{k-1})$ for multinomial distribution is given as:

$$g(p_1 \cdots p_{k-1}) = \left(\prod_{j=1}^{k-1} p_j^{s_j - 1} \right) \left(1 - \sum_{j=1}^{k-1} p_j \right)^{s_k - 1} \qquad (16)$$

(see Good (1965). For experimental outcome $(n_1, n_2 \ldots, n_k)$ the likelihood function is given by:

$$\ell(n_1 \cdots n_k) = w(n_1, n_2 \cdots n_k) \left(\prod_{j=1}^{k-1} p_j^{n_j} \right) \left(1 - \sum_{j=1}^{k-1} p_j \right)^{n_k} \qquad (17)$$

where $w(n_1, n_2, \ldots, n_k)$ is the total number of admissible paths from $(0, 0, \ldots, 0)$ to $(n_1, n_2, \ldots, n_k)$. (Note: $w(n_1, n_2, \ldots, n_k) \neq \left(\frac{m!}{\prod_{j=1}^{k} n_j!} \right)$, since for sequential procedures the outcomes are not permutable. Fortunately, however, $w(n_1, n_2, \ldots, n_k)$ does not present any problem since it drops out, as may be seen later.)

Substituting values of $g(p_1 \cdots p_{k-1})$ and $\ell(n_1 \cdots n_k)$ in (15) the posterior density becomes:

$$f(p_1 \cdots p_{k-1}) = \frac{\left(\prod_{j=1}^{k-1} p_j^{r_j}\right)\left(1-\sum_{j=1}^{k-1} p_j\right)^{r_k}\left(\prod_{j=1}^{k-1} dp_j\right)}{\int_0^1 \int_0^{1-p_1} \cdots \int_0^{1-\sum_{j=1}^{k-1} p_j} \left(\prod_{j=1}^{k-1} p_j^{r_j}\right)\left(1-\sum_{j=1}^{k-1} p_j\right)^{r_k}\left(\prod_{j=1}^{k-1} dp_j\right)} \tag{18}$$

$$= \frac{\left(\prod_{j=1}^{k-1} p_j^{r_j}\right)\left(1-\sum_{j=1}^{k-1} p_j\right)^{r_k}\left(\prod_{j=1}^{k-1} dp_j\right)}{B(r_1+1,\ r_2+1,\ldots,r_k+1)} \tag{19}$$

where $r_j = n_j + s_j - 1; \quad \forall j$

and

$$B(r_1+1,\ r_2+1,\ldots,r_{k+1}) = \frac{\left(\prod_{j=1}^{k} r_j!\right)}{(\Sigma(r_j+1)-1)!}$$

(see Appendix 1)

Let $\hat{p}_i$ be the point estimate of p_i; $\hat{p} = E(p)$. It follows that

$$\hat{p}_i = \frac{\int_0^1 \int_0^{1-p_1} \cdots \int_0^{1-\sum_{j=1}^{k-1} p_j} p_i \left(\prod_{j=1}^{k-1} p_j^{r_j}\right)\left(1-\sum_{j=1}^{k-1} p_k\right)^{r_k}\left(\prod_{j=1}^{k-1} dp_j\right)}{B(r_1+1,\ r_2+1,\ldots r_k+1)}$$

$$= \frac{B(r_1+1,\ldots,r_{i-1}+1,\ r_i+2,\ r_{i+1}+1,\ldots,r_k+1)}{B(r_1+1,\ r_2+1,\ldots r_k+1)}$$

$$= \frac{r_i+1}{((\Sigma r_i)+k)}$$

$$= \frac{n_i+s_i}{m+\sum_{j=1}^{k} s_j} \tag{20}$$

3.2 Interval estimation

Let $\underline{p}_i$ be the lower bound on p_i at confidence level $1-\alpha$ (i.e., $\Pr(p_i \le \underline{p}_i) \le \alpha$). Without loss of generality we may interchange p_i with p_1 and relabel these two. It follows that

$$\frac{\int_0^{p_1} \int_0^{1-p_1} \cdots \int_0^{1-\sum_{j=1}^{k-1} p_j} \left(\prod_{j=1}^{k-1} p_j^{r_j}\right)\left(1-\sum_{j=1}^{k-1} p_j\right)^{r_k} \left(\prod_{j=1}^{k-1} dp_j\right)}{B(r_1+1,\ r_2+1,\ldots,\ r_k+1)} = \alpha$$

$$\text{or} \quad \frac{B_{\underline{p}}\left(r_1+1,\ \left(\sum_{j=2}^{k} r_j\right)+k-1\right)}{B\left(r_1+1,\left(\sum_{j=2}^{k} r_j\right)+k-1\right)} = \alpha \qquad (21)$$

$$\text{or} \quad E\left(\sum_{j=1}^{k} r_j+k-1,\ r_1+1,\ \underline{p}_1\right) = \alpha \qquad (22)$$

Thus, in general, $\underline{p}_i$ may be obtained by solving one of the following two equations.

$$\frac{B_{\underline{p}_i}\left(n_i+s_i,\ m-n_i+\sum_{\substack{j=1\\ j\neq i}}^{k} s_j\right)}{B\left(n_i+s_i,\ m-n_i+\sum_{\substack{j=1\\ j\neq i}}^{k} s_j\right)} = \alpha \qquad (23)$$

$$\text{or} \quad E\left(m+\sum_{j=1}^{k} s_j-1,\ n_i+s_i,\ \underline{p}_i\right) = \alpha \qquad (24)$$

Similarly, an upper bound $\bar{p}_i$ on p_i at confidence level $\beta (\beta \le 1)$; (i.e. $\Pr(p_i \ge \bar{p}_i) \le \beta$ or $\Pr(p_i \le \bar{p}_i) \ge 1-\beta$) may be computed by

solving either (23) or (24) by changing their r.h.s. to $1-\beta$.

The two-sided confidence interval at confidence level $1-\alpha-\beta$ ($\alpha,\beta<1$; α from the lower side and β from the upper side) may be obtained by solving (21) or (22) twice--first with α and next with $1-\beta$ in r.h.s.

3.3 Results with uniform prior.

For uniform prior distribution of $(p_1 \ldots p_k)$, p_j between 0 and 1 (inclusive) $g(p_1 \ldots p_{k-1})=1$; i.e. $s_j=1$ (for all j) in (16). Substituting above values of s_j in (20)

$$\hat{p}_i = \frac{n_i+1}{m+k} \tag{24}$$

The lower bound $\underline{p}_i$ on p_i at confidence level α may be obtained by solving

$$B_{\underline{p}_i}(n_i+1,\ m-n_i+k-1)/B(n_i+1,\ m-n_i+k-1) = \alpha \tag{25}$$

or $$E(m+k-1,\ n_i+1,\ \underline{p}_i) = \alpha \tag{26}$$

Similarly, the upper bound $\bar{p}_i$ on p_i at confidence level β may be obtained by solving one of the following two equations:

$$B_{\bar{p}}(n_i+1,\ m-n_i+k-1)/B(n_i+1,\ m-n_i+k-1) = 1-\beta \tag{27}$$

$$E(m+k-1,\ n_i+1,\ \bar{p}_i) = 1-\beta \tag{28}$$

The two-sided bound $(\underline{p}_i \bar{p}_i)$ on p_i at confidence level $\alpha+\beta$ ($\alpha+\beta<1$) may be obtained by solving (25) and (27) or (26) and (28) simultaneously.

3.4 Illustrative example

As an illustration of the explicit point and interval estimation, suppose our sequential experiment for selecting the most probable event of a multinomial experiment (of order 3; k=3) terminates at stage 7 with score [4,2,1]. It is desired to obtain

(a) Point estimates of p_1, p_2 and p_3

(b) One-sided bound on p_1, p_2, and p_3 at confidence level 0.9

(c) Two-sided bound on p_1, p_2, and p_3 at confidence level 0.9.

Solution

(a) From (24)

$$\hat{p}_1 = \frac{4+1}{7+3} = 0.5,\ \hat{p}_2 = \frac{2+1}{7+3} = 0.3,\ \hat{p}_3 = \frac{i+1}{7+3} = 0.2$$

(b) (i) Using (25) with $\alpha = 0.1$, the lower bound (at 90%)

$$\underline{p}_1 = 0.300,\ \underline{p}_2 = 0.129,\ \underline{p}_3 = 0.060$$

(ii) Using the same equation with $\beta = 0.9$ (in place of α, the upper bound at 90%

$$\overline{p}_1 = 0.699,\ \overline{p}_2 = 0.49,\ \overline{p}_3 = 0.368$$

(c) Using (25) first with 0.05 and then 0.95 in r.h.s. the 90% two-sided bounds on p_i's are:

$$0.251 \leq p_1 \leq 0.748$$

$$0.097 \leq p_2 \leq 0.549$$

$$0.041 \leq p_3 \leq 0.429$$

REFERENCES

1. Anscombe, F.J., (1953),"Sequential Estimation (with discussion)", J.Roy.Statist.Soc., B, Vol. 15, pp.1-29.
2. Anscombe, F.J., (1954), "Fixed-Sample Size Analysis of Sequential Observations", Biometrics, Vol. 10, pp. 89-100.
3. Bhat, B.R. and Kulkarni, N.V. (1966), "On Efficient Multinomial Estimation:, J.Roy.Statist. Soc., Vol. 28, pp. 45-52.
4. Box, G.E.P. and Tiao, G.C. (1973), Bayesian Inference in Statistical Analysis, Addison-Wesley, Reading, Mass.
5. Chandra, R. (1975-A), Truncated Sequential Procedures for Discriminating among k≥2 Hypotheses Concerning Binomial Parameter p, AES Monograph 7511, Union College, Schenectady, N.Y.
6. Chandra, R. (1975-B), Truncated Sequential Procedures for the Selection of the Most Probable Category of a Multinomial Population, AES Monograph 7512, Union College, Schenectady, N.Y.
7. Chandra, R. (1975-C), Truncated Sequential Procedures for Complete Ranking of Several Binomial Populations, AES Monograph 7513, Union College,Schenectady, N.Y.
8. DeGroot, M.H. (1959), "Unbiased Sequential Estimation for Binomial Population", Ann. Math. Statist., Vol. 30, pp. 80-101.

9. DeGroot, M.H. (1970), Optimum Statistical Decisions, McGraw-Hill, Hightstown, N.J.

10. Girshick, M.A., Mosteller, F. and Savage, L.J. (1946), "On Biased Estimates for Certain Binomial Sampling Problems with Applications", Ann. Math. Statist., Vol. 17, pp. 13-23.

11. Good, I.J. (1965), The Estimation of Probabilities, M.I.T. Press, Cambridge, Mass.

12. Good, I.J. and Crook, J.F. (1974), "The Bayes/Non-Bayes Compromises and the Multinomial Distribution", J.Amer. Statist. Assoc., Vol. 69, pp. 711-720.

13. Goss, T.I. (1974), Nonparametric Truncated Sequential Tests for the Median with Application to Maintainability, AES Monograph 745, Union College, Schenectady, N.Y.

14. Harvard University Computation Laboratory (1955), Table of Cumulative Binomial Probability Distribution, Harvard University Press, Cambridge, Mass.

15. Johnson, N.L. (1961), "Sequential Analysis: A Survey", J. Of Royal Statist. Soc., A, Vol. 124, pp. 372-411.

16. Johnson, N.L. and Kotz, S. (1969), Distributions in Statistics: Discrete Distributions, Houghton Mifflin, Boston, Mass.

17. Johnson, W.E. (1932), Appendix (ed. by R.B. Braithwaite) to "Probability Deductive and Inductive Problems", Mind, Vol. 41, 421-423.

18. Muhamedhanova, R. and Suleimanova M. (1961), "Unbiased Estimates for Selection Problems in the Case of Multinomial Distribution", Proc. of Inst. of Math., AN UZSSR, 22, pp. 121-129.

19. Perk, W. (1947), "Some Observations on Inverse Probability Including a New Indifference Rule", J. Inst. Actuar., Vol. 73, pp. 285-312.

APPENDIX 1

SOME DERIVATIONS RELATED WITH MULTINOMIAL ESTIMATION

1. Derivation of (19).

In what follows we prove that

$$I_k = \int_0^1 \int_0^{1-p_1} \cdots \int_0^{1-\sum_{j=1}^{k-1} p_j} \left(\prod_{j=1}^{k-1} p_j^{r_j} \right) \left(1-\sum_{j=1}^{k-1} p_j \right)^{r_k} \left(\prod_{j=1}^{k-1} dp_j \right)$$

$$= B(r_1+1,\ r_2+r_3+\ldots+r_k+k-1)\ B(r_2+1,\ r_3+\ldots+r_k+k-2)$$
$$\ldots\ B(r_i+1,\ r_2+1,\ \ldots,\ r_k+1)$$

$$= B(r_i+1,\ r_2+1,\ \ldots,\ r_k+1)$$

(i) For k=2

$$I_2 = \int_0^1 p_1^{r_1} (1-p_1)^{r_2}\, dp_1$$

$$= B(r_1+1,\ r_2+1) \qquad \text{(A-1)}$$

(ii) For k=3

$$I_3 = \int_0^1 \int_0^{1-p_1} p_1^{r_1} p_2^{r_2} (1-p_1-p_2)^{r_3} dp_2 dp_1$$

Let $\frac{p_2}{1-p_1} = x$, substituting in r.h.s.

$p_2 = (1-p_1)x$, $dp_2 = (1-p_1)dx$ and changing the limits of integration from $0-(1-p_1)$ to 0-1 we have

$$I_3 = \int_0^1 \int_0^1 p_1^{r_1} x^{r_2} (1-p_1)^{r_2} (1-p_1)^{r_3} (1-x)^{r_3} (1-x)\,dx\,dp_1$$

$$= \int_0^1 p_1^{n_1} (1-p_1)^{r_2+r_3+1} \int_0^1 x^{r_2} (1-x)^{r_3}$$

$$= B(k_1+1\ r_2+r_3+2)\ B(r_2+1,\ r_3+1)$$

$$= \frac{r_i!\ (r_2+r_3+1)!}{(r_1+1+r_2+r_3+2-1)} \frac{r_2!\ r_3!}{(r_2+1+r_3+1-1)} \qquad \text{(A-2)}$$

$$= \frac{r_1!\ r_2!\ r_3!}{(r_1+r_2+r_3+3+1)}$$

$$= B(r_1+1,\ r_2+r_3+1)$$

Thus following through we may generalize from (A-1) and (A-2) that

$$I_k = B(r_1+1,\ r_2+\ldots+r_k+k-1)\ B(r_2+1,\ R_3+\ldots+r_k+k-2)$$
$$B(r_{k-1}+1,\ r_k+1),$$

From (A-1) and (A-2) it follows that

$$I_k = B(r_1+1,\ r_2+1,\ldots,r_k+1)$$

Q.E.D.